

UNIVERSIDAD PANAMERICANA

Probabilidad de ganar o perder al invertir en acciones de la BMV

Mayo 2020

Econometría II

Manuel Alejandro Díaz Rodríguez

Eliut Rodríguez Bautista.

Resumen ejecutivo.

El presente trabajo presenta un modelo de regresión logística aplicado al ámbito financiero, de tal forma que se pueda determinar la probabilidad de obtener ganancias invirtiendo en acciones pertenecientes a la Bolsa Mexicana de Valores. En la primera parte del documento se muestra la motivación que nos llevó a realizar la presente investigación y los objetivos que buscamos alcanzar a lo largo de esta; posteriormente se presentan algunas ideas, conceptos y modelos que sirvieron como base en la definición de variables. En seguida se eligen las empresas objetivo, el tamaño de muestra y el proceso que se siguió para que todas las variables tuvieran la misma amplitud, después se muestran algunas evoluciones temporales de los rendimientos, cuya característica más observable es la volatilidad.

Pasando a la construcción del modelo hacemos énfasis en la periodicidad de los datos y el problema más importante generado por dicha decisión, el ruido innecesario, el cual se materializa en la volatilidad anteriormente citada. En la estimación se muestran los estimadores β para cada variable e inmediatamente nos ocupamos en evaluar la capacidad predictiva del modelo utilizando dos tipos de límites, el arbitrario 0.5 y el ratio de éxitos observados, concluyendo que las predicciones son en su mayoría buenas. Posteriormente se muestran los diversos efectos producidos por cambios marginales en variables catalogadas como de interés, dada su relación a priori con el objetivo del trabajo (precio de la acción) y con el giro de las empresas elegidas, se puede apreciar un fuerte efecto ante cambios en la tasa del bono mexicano a un mes y en el tipo de cambio. Finalmente se muestran las conclusiones, con ello se exponen algunas extensiones y posibles cambios en la definición de variables y empresas a revisar.

Objetivo General.

- El objetivo del presente trabajo es determinar la probabilidad promedio de que un inversionista obtenga ganancias de capital al invertir en la Bolsa Mexicana de Valores, aplicando el modelo de Regresión Logística a una serie de variables de corte económico-financiero

Objetivos Particulares.

- Aplicar el modelo de regresión logística a cuatro empresas consideradas como representativas de la Bolsa Mexicana de Valores.
- Aprender a usar distintas herramientas para la obtención de información útil y posterior construcción de bases de datos.
- Implementar el uso del lenguaje de programación “Python” en la modelación y análisis econométrico.

Introducción.

- *¿Cuál es el problema y motivación del trabajo?*

Como bien sabemos el ambiente de inversión y, en general, el panorama económico no es nada alentador. En la misma línea, hay un consenso generalmente aceptado el cual establece que todos los modelos colapsan o explican erróneamente la realidad cuando pasamos por una crisis. Sin embargo, creemos que esta situación es temporal y, a pesar de que la recuperación lleve mucho tiempo, hay una clara ventana de oportunidad. Para reactivar la economía se necesita certidumbre,

seguridad de que se obtendrán ganancias, o bien, que no se perderá a la hora de invertir. De esta forma concluimos que se necesita una metodología o modelo que determine la probabilidad promedio de ganar o perder en el mercado bursátil.

Las inversiones siempre han sido una gran fuente de ingresos, históricamente han sido el motor del crecimiento en todas las economías; por medio de distintos mecanismos y fuentes de financiamiento, las empresas, gobiernos y familias pueden innovar, sacar nuevos productos, incrementar la producción y adquirir activos. Actualmente los mercados donde se lleva a cabo dicha actividad buscan incrementar el valor de las empresas y retribuir a los participantes que financien a las mismas. La idea de transferir capitales se fundamenta en el hecho de que los participantes buscan maximizar sus beneficios. Este problema de optimización aparte de considerar al ingreso como una restricción, también necesita tomar en cuenta un factor clave, el riesgo, que para efecto del presente trabajo se definirá como la variación en los precios.

Con base a lo anteriormente expuesto, la motivación del presente trabajo es crear un modelo, que aplique las herramienta e ideas adquiridas a lo largo de la carrera, que le permita a cualquier inversionista potencial o veterano saber si sus inversiones, en nuestro país y más en específico en la Bolsa Mexicana de Valores, darán o no frutos.

- *Literatura.*

Existen modelos que sirven para determinar la rentabilidad de los activos financieros con el fin de asignarlos a un portafolio o para invertir directamente en ellos. El Capital Asset Pricing Model (CAPM) es muy útil para fines de inversión, al representar una de las mejores alternativas para el cálculo de la tasa de retorno exigida por los inversionistas (Ross, Westerfield & Jaffe, 1995). Este método fue desarrollado por William F. Sharpe (1964), John Lintner (1965) y Jan Mossin (1966), en base a los trabajos de Harry Markowitz. Para ellos los inversionistas optan por aquellas inversiones que dan el mayor rendimiento esperado para un determinado nivel de riesgo, considerando que siempre hay una opción de inversión libre de riesgo. El CAPM puede estimarse con el siguiente modelo:

$$E(r_i) = r_{f0} + \beta * (E(r_m) - r_{f0})$$

Donde:

$E(r_i)$ = Rendimiento esperado del activo o inversión.

R_{f0} = Tasa libre de riesgo.

$E(r_m)$ = Tasa de rendimiento promedio de los activos disponibles en el mercado.

Otro modelo que también sirve para la estimación de los rendimientos de un activo es la Teoría de Valoración por Arbitraje o APT, propuesta por Stephen A. Ross en 1976. Para este autor el retorno esperado de los activos con riesgo depende de una combinación lineal de K factores, es decir, la utilidad de un activo se representa como una función lineal de K elementos.

De manera básica la APT considera un retorno mínimo en ausencia de riesgo (tasa libre de riesgo), la tasa de desempleo para cada periodo estudiado, el índice de la Bolsa a estudiar (símil del rendimiento esperado de mercado en el modelo CAPM), la tasa de inflación (medida con el INPC de los periodos a estudiar) y el riesgo país de la bolsa a estudiar. Como se puede observar el rendimiento de un activo es influenciado por el riesgo sistemático, es decir, por las fluctuaciones en

variables macroeconómicas no controladas por el público inversor. Por otro lado, está el famoso de Teorema de Modigliani Miller que, en su segunda proposición en un mercado de capitales imperfecto (cuando hay impuestos corporativos y no hay una libre entrada y salida de oferentes y demandantes), se puede determinar el costo del capital accionario, es decir, la tasa de rendimiento que piden los inversionistas para invertir en una empresa, tomando como variables explicativas a la tasa fiscal, el ratio deuda-capital y el costo de la deuda (tasa de interés que cobran los acreedores o prestamistas).

Explicación detallada del modelo.

Conceptos relevantes:

- Ganancias de capital.

Según el Diccionario Jurídico Mexicano, son los incrementos en el valor de realización respecto al de adquisición del patrimonio poseído, siendo estas inesperadas o esporádicas.

- Riesgo de financiero.

Según el Banco BBVA, este concepto hace referencia a las posibles variaciones en el rendimiento de una inversión, derivado de los cambios en el sector al que pertenece la emisora, la inestabilidad de los mercados y la imposibilidad de devolución del capital por una de las partes.

- Split o desdoblamiento de acciones.

Es el aumento del número de acciones, sin que el capital social de la empresa varíe. Esto implica que el valor nominal y la cotización en el mercado bursátil de un activo sea menor, dada una proporción establecida.

- Nivel de capitalización bursátil.

También conocido como valor en bolsa, es aquella medida económica que indica el valor total de todas las acciones cotizadas de una empresa; esta medida se obtiene multiplicando el número de acciones que hay en circulación por el precio de mercado.

Establecimiento del modelo.

El modelo trata de integrar las ideas expuestas anteriormente. Para modelar la probabilidad de ganar o perder al momento de adquirir una acción en la BMV, se buscaron dos tipos de variables. Las primeras buscan explicar la magnitud de las ganancias o pérdidas reportadas por las emisoras, a lo largo del tiempo. La segunda categoría de variables se eligieron con el afán de medir las fuentes de riesgo más importantes, dado los sectores a los que pertenecen las emisoras.

Las variables a elegir fueron:

1. *Los precios de cierre diarios de cada emisora:* Perteneciente a la primera categoría, es la principal fuente de información a la hora de calcular ganancias o pérdidas de capital.
2. *El volumen de acciones en circulación diarios de cada emisora:* Perteneciente a la primera categoría, un Split de acciones normalmente lleva a una caída en el precio de las acciones, es decir, existe una relación inversa entre el número de acciones y el precio de las mismas

3. *El índice de Precios y Cotizaciones*: Perteneciente a la primera categoría, esta variable servirá como proxy al rendimiento de mercado, este índice sirve como referencia a la hora de comparar acciones y portafolios en específico, si el rendimiento particular es menor al rendimiento reportado por el IPC, generalmente no se invierte en ese activo.
4. *El volumen del Índice de Precios y Cotizaciones*: la idea es que funcione como un símil al volumen de acciones en circulación de cada emisora de forma agregada o para todo el mercado.
5. *Tipo de Cambio (USD/MXN)*: Perteneciente a la segunda categoría, consideramos esta variable como importante, porque las emisoras elegidas tienen gran dependencia al sector extranjero. Cambios en esta variable pueden representar pérdidas o beneficios en las empresas y posteriormente incrementos los rendimientos reportados por las firmas.
6. *La tasa del Bono a 1 mes de México*: Perteneciente a la segunda categoría, es equivalente a la tasa libre de riesgo, la cual contempla en su determinación a la tasa de inflación y el riesgo país.
7. *El rendimiento de cada acción*: Para crear la variable “Y” en su forma dicotómica, se obtuvo el rendimiento por acción, que para efectos del presente trabajo se definirá como la variación diaria de los precios de cada título. El valor 0 se le dio a aquellos rendimientos con signo negativo y 1 a los que presentaron signo positivo.

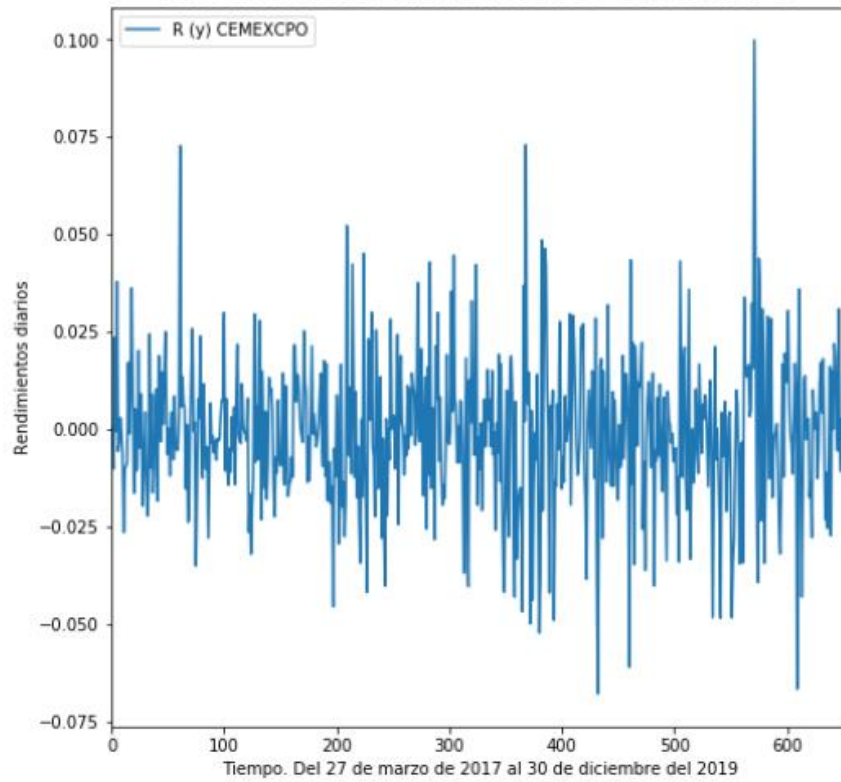
Las cuatro emisoras se eligieron en función de su nivel de capitalización bursátil siendo estas: Grupo México, Cemex, América Móvil y Bimbo.

Definición de muestra y balanceo de base de datos.

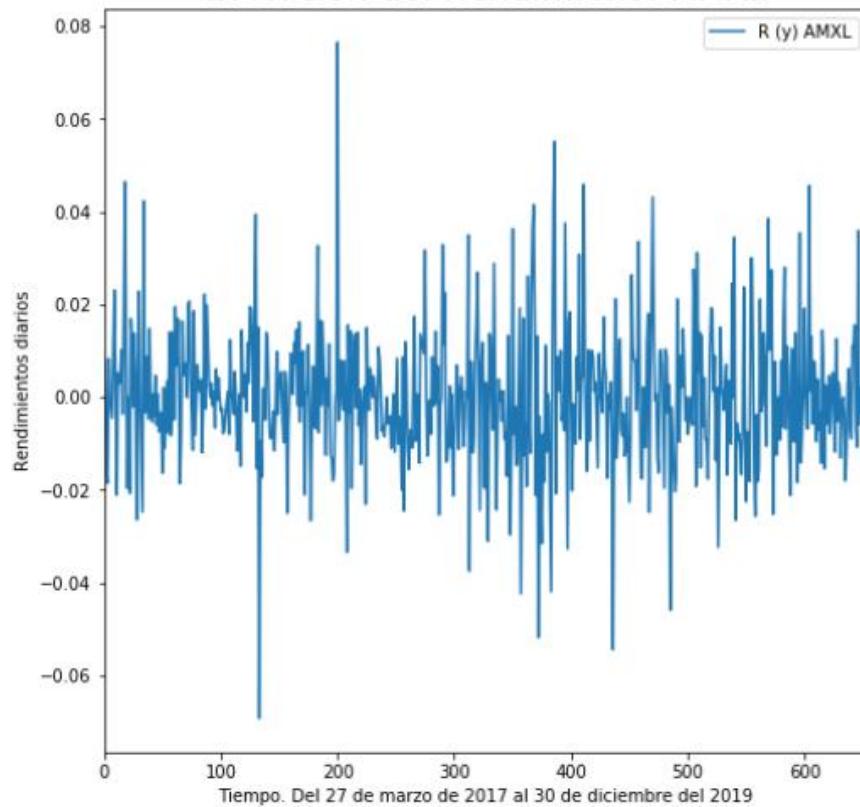
Con el fin de observar un comportamiento más real de las variables, se optó por una periodicidad diaria. Siendo el periodo a estudiar el comprendido desde el primero de marzo de 2017 al 31 de diciembre de 2019. Tal como se explicó en la introducción, la función predicativa o explicativa del modelo falla en momentos de crisis y como el 2020 inició con una, creímos que era buena idea estudiar hasta 2019. Por otro lado, el número de observaciones era diferente entre variables, el número de datos del volumen y precio tanto del índice como de las acciones era menor al número de observaciones en el tipo de cambio y la tasa del bono mexicano a 30 días. Para resolver este problema, se identificaron y, posteriormente se eliminaron, los días donde había ausencia de información, los cuáles concordaban con fines de semana y algunos días festivos, es decir, cuando no hay cotización u operación en la BMV.

Finalmente, la muestra se conformó de 656 observaciones para cada una de las 6 variables anteriormente expuestas. A continuación se presentan algunos ejemplos de la evolución histórica de las variables.

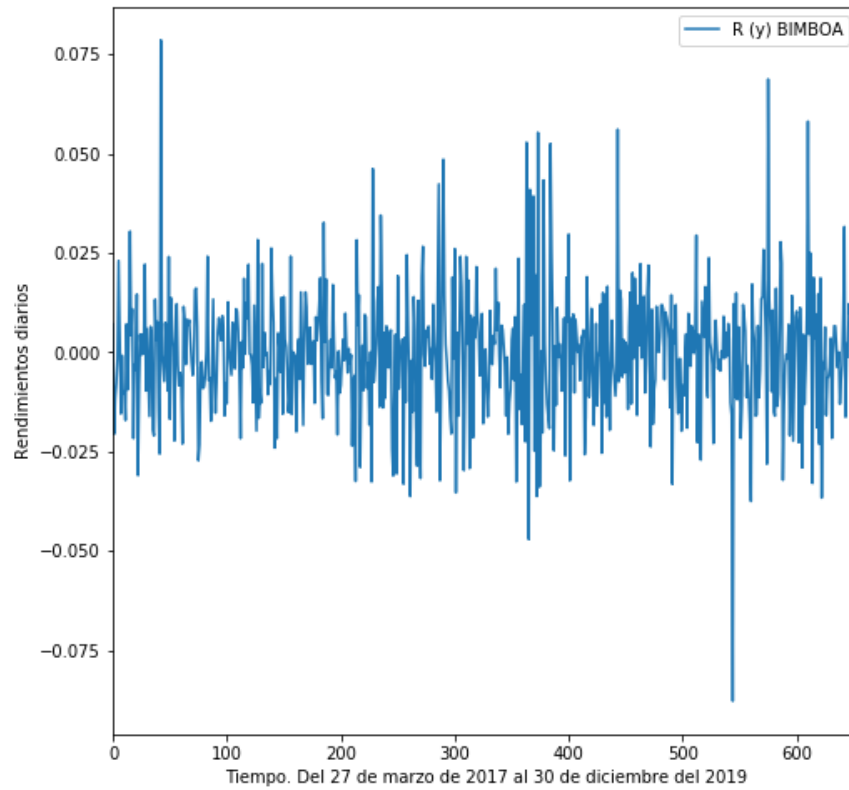
Evolución del Rendimiento CEMEXCPO



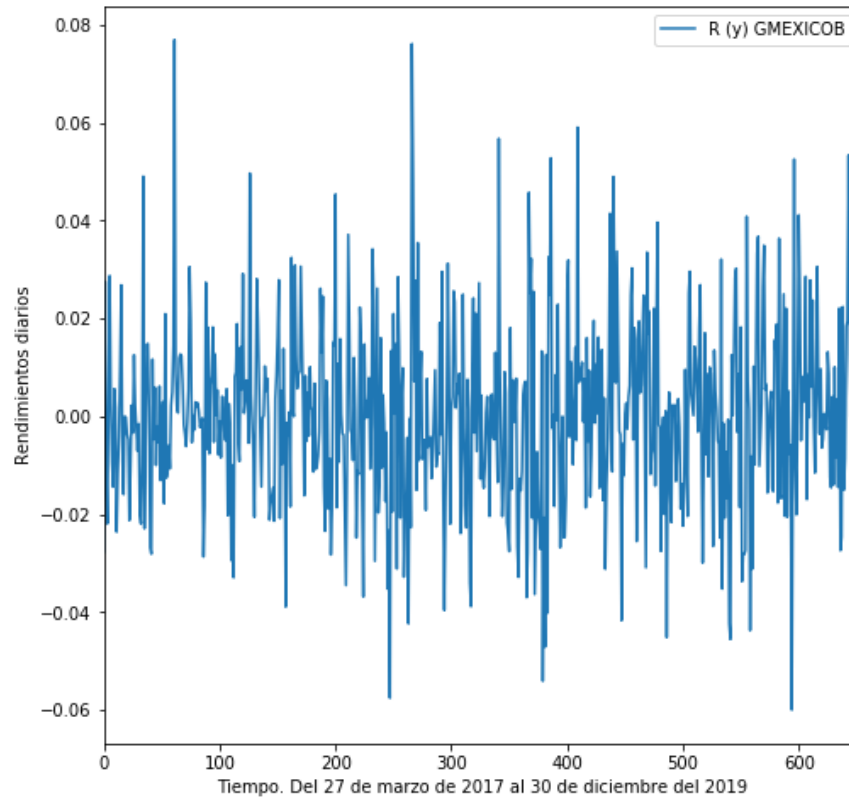
Evolución del Rendimiento AMXL



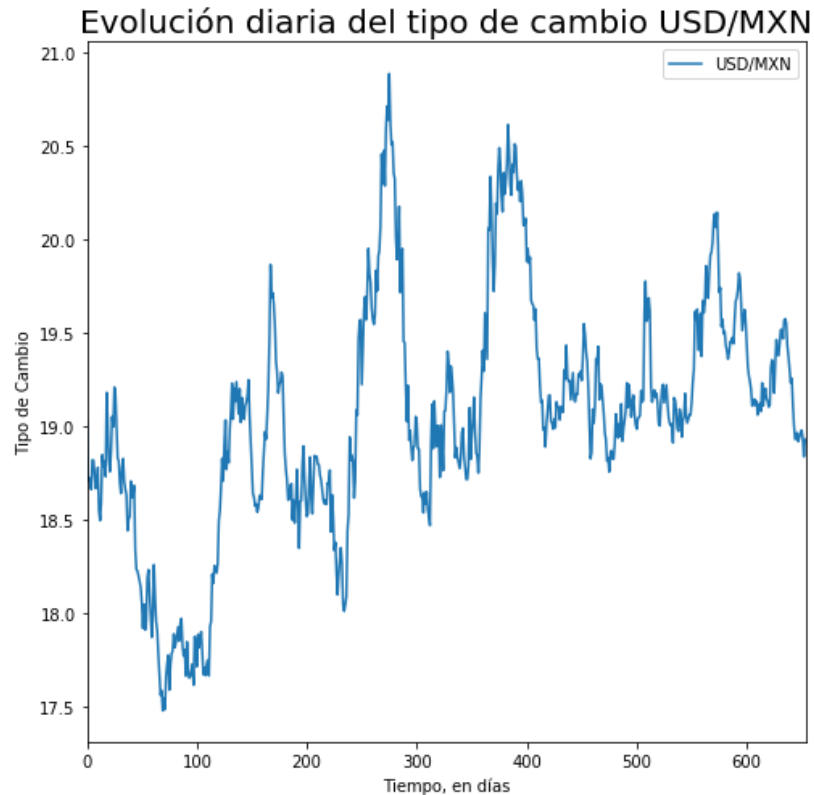
Evolución del Rendimiento BIMBOA



Evolución del Rendimiento GMEXICOB



Los gráficos anteriores muestran una gran variabilidad en los rendimientos diarios, cuyas medias, aunque tienden a 0, no son capaces de representar al universo de datos. Esto puede deberse a la gran sensibilidad (que en el modelo CAPM se mide con la β) de las empresas respecto del mercado. Generalmente, las coyunturas de muy corta duración agregan ruido (riesgo) innecesario a la hora de invertir.



Por su parte el tipo de cambio, a modo de ejemplo presenta una media de 19.0469 pesos por dólar con una variabilidad de 64.09 centavos, un valor mínimo de 17.4790 y un máximo de 20.8905. La gráfica, por su lado, muestra una pequeña tendencia a la alza, con varios picos importantes.

Especificación del modelo.

Tal como se expuso en el objetivo general se usará el modelo logístico:

$$Prob(Y = 1) = \frac{1}{1 + e^{-z}}; \text{ donde } z = \alpha + \beta_k x_{ki} + u_i$$

Ya definidas las variables regresoras (x_{ki}), la variable de respuesta (y_i) y su comportamiento a lo largo del tiempo. La regresión z tiene la siguiente forma:

$$y_t = \beta_0 + \beta_1 Close_t + \beta_2 Volume_t + \beta_3 IPC\ close_t + \beta_4 IPC\ volume_t + \beta_5 TC\ close_t + \beta_6 MX\ 1M\ close_t + u_t$$

Donde:

- Y_t = Es una variable dicotómica con valores 1 y 0, dependiendo del signo del rendimiento diario.

- $Close_t$ = Precios de cierre diario de cada empresa.
- $Volume_t$ = Volumen de acciones en circulación diarios de cada empresa.
- $IPC\ close_t$ = Índice de Precios y Cotizaciones diario.
- $IPC\ volume_t$ = Volumen diario del IPC.
- $TC\ close_t$ = Tipo de cambio USD/MEX diario.
- $MX\ 1M_t$ = Tasa de interés del bono mexicano de 30 días.

Estimación del modelo.

Con el lenguaje de programación Python se realizaron las 4 regresiones logísticas, las cuales arrojaron los siguientes resultados:

Regresión para CEMEXCPO:

Dep. Variable:	y	No. Observations:	656
Model:	Logit	Df Residuals:	650
Method:	MLE	Df Model:	5
Date:	Thu, 14 May 2020	Pseudo R-squ.:	0.008181
Time:	10:44:35	Log-Likelihood:	-447.27
converged:	True	LL-Null:	-450.96
Covariance Type:	nonrobust	LLR p-value:	0.1940

	coef	std err	z	P> z	[0.025	0.975]
CEMEXCPO_Close	0.0332	0.059	0.559	0.576	-0.083	0.149
CEMEXCPO_Volume	3.051e-09	6.82e-09	0.447	0.655	-1.03e-08	1.64e-08
IPC_Close	-4.012e-05	4.58e-05	-0.876	0.381	-0.000	4.96e-05
IPC_Volume	1.624e-09	1.72e-09	0.945	0.345	-1.74e-09	4.99e-09
USD/MXN_Close	0.0793	0.095	0.834	0.404	-0.107	0.266
MX 1M_Close	-0.0845	0.177	-0.478	0.633	-0.431	0.262

Regresión para BIMBO A:

Dep. Variable:	y	No. Observations:	656
Model:	Logit	Df Residuals:	650
Method:	MLE	Df Model:	5
Date:	Thu, 14 May 2020	Pseudo R-squ.:	0.006336
Time:	10:44:35	Log-Likelihood:	-451.39
converged:	True	LL-Null:	-454.27
Covariance Type:	nonrobust	LLR p-value:	0.3306

	coef	std err	z	P> z 	[0.025	0.975]
BIMBOA_Close	0.0536	0.032	1.694	0.090	-0.008	0.116
BIMBOA_Volume	8.395e-08	5.56e-08	1.509	0.131	-2.51e-08	1.93e-07
IPC_Close	-4.088e-05	3.36e-05	-1.219	0.223	-0.000	2.49e-05
IPC_Volume	-9.872e-10	1.37e-09	-0.719	0.472	-3.68e-09	1.7e-09
USD/MXN_Close	-0.0360	0.087	-0.413	0.679	-0.207	0.135
MX 1M_Close	0.0392	0.160	0.245	0.806	-0.274	0.352

Regresión para AMXL:

Dep. Variable:	y	No. Observations:	656
Model:	Logit	Df Residuals:	650
Method:	MLE	Df Model:	5
Date:	Thu, 14 May 2020	Pseudo R-squ.:	0.008591
Time:	10:44:35	Log-Likelihood:	-447.09
converged:	True	LL-Null:	-450.96
Covariance Type:	nonrobust	LLR p-value:	0.1707

	coef	std err	z	P> z 	[0.025	0.975]
AMXL_Close	0.0139	0.083	0.168	0.866	-0.148	0.176
AMXL_Volume	-5.99e-09	6.67e-09	-0.898	0.369	-1.91e-08	7.09e-09
IPC_Close	-2.031e-05	2.88e-05	-0.705	0.481	-7.67e-05	3.61e-05
IPC_Volume	4.085e-09	2.26e-09	1.808	0.071	-3.44e-10	8.51e-09
USD/MXN_Close	0.0564	0.090	0.630	0.528	-0.119	0.232
MX 1M_Close	-0.1314	0.165	-0.798	0.425	-0.454	0.191

Regresión para GMEXICOB:

Dep. Variable:	y	No. Observations:	656
Model:	Logit	Df Residuals:	650
Method:	MLE	Df Model:	5
Date:	Thu, 14 May 2020	Pseudo R-squ.:	0.009060
Time:	10:44:35	Log-Likelihood:	-446.88
converged:	True	LL-Null:	-450.96
Covariance Type:	nonrobust	LLR p-value:	0.1471

	coef	std err	z	P> z	[0.025	0.975]
GMEXICOB_Close	-0.0138	0.018	-0.759	0.448	-0.049	0.022
GMEXICOB_Volume	1.509e-08	2.16e-08	0.699	0.484	-2.72e-08	5.74e-08
IPC_Close	1.082e-06	3.14e-05	0.034	0.973	-6.05e-05	6.26e-05
IPC_Volume	1.651e-09	1.39e-09	1.189	0.234	-1.07e-09	4.37e-09
USD/MXN_Close	0.0490	0.089	0.552	0.581	-0.125	0.223
MX 1M_Close	-0.1141	0.162	-0.703	0.482	-0.432	0.204

Presentación e interpretación de resultados.

Como bien sabemos este tipo de modelos sólo nos proporciona resultados entre 0 y 1. Por tanto, debemos elegir cómo se han de clasificar dichos resultados, es decir, definir un límite (threshold) que determine cuáles serán 1 (que se gane en bolsa) o 0 (que se pierda en bolsa).

El presente trabajo usará límites de 0.5 y el resultado del ratio de éxitos que por definición es el cociente de éxitos observados entre el total de observaciones. El primero se eligió porque en los 4 casos la media es muy cercana a 0, es decir, en el periodo estudiado se gana y se pierde, en promedio, en la misma proporción. Del mismo modo se eligió al ratio de éxitos ya que este se ajusta más a la realidad (datos observados), y aunque este método pueda sobreestimar el número de casos de éxitos explicados.

En las cuatro regresiones, la capacidad predictiva de éxitos no varió mucho entre los límites usados. Por ejemplo, el modelo para AMXL con un límite de 0.5, predijo 61 casos de éxito de los 314 observados, siendo correctos un 51.908% de estos; con el ratio de éxitos, se predijeron 261 casos de éxito de los 314, siendo correctos en el 52.213% de los casos.

Para CEMEXCPO, con un límite de 0.5 se predijeron 64 casos de 292, siendo correctos en un 57.251%; por otro lado con el ratio de éxitos, se logró predecir 255 casos de 292, con un porcentaje de efectividad de 55.572%. El caso de BIMBOA con un límite de 0.5 se predijo 220 casos de 315 con una capacidad predictiva de 54.045%; con el ratio de éxitos se sobreestimaron el número de éxitos, siendo estos un total de 326, con una probabilidad de acertar de 51.908%. Para finalizar GMEXICOB,

con un límite de 0.5 se predijo 72 casos de 319, con una efectividad de 49.77%; el ratio de éxito logró pronosticar 262 casos con una probabilidad acertar de 50.687%.

Por otro lado, utilizando los conceptos PEA y APE se pudieron encontrar los efectos que podían suceder en la probabilidad de ganar si modificáramos marginalmente unas variables consideradas de interés, entre las que figuran: el precio de cierre diario de la acción, el tipo de cambio y la tasa del bono. Recordando que este efecto, dada la naturaleza no lineal del modelo, no puede ser el valor del estimador β correspondiente a la variable.

Para AMXL, considerando el PEA, el cambio en un peso del precio de cierre de la acción ocasionará un cambio en el promedio de la probabilidad de ganar en la bolsa de aproximadamente 0.621%; considerando el APE, la probabilidad de ganar en la bolsa se incrementa o decrece en -6.962%, si el tipo de cambio se aprecia o deprecia en 1 peso. En el caso de CEMEXCPO, la depreciación o apreciación del tipo de cambio en un peso generará un cambio de 3.539% en la probabilidad de ganar o perder en este título, usando el PEA. Por otro lado y usando el concepto APE, un cambio en el precio de cierre generará una caída o aumento en la probabilidad de ganar de 4.227%.

Específicamente en GMEXICOB, un cambio en la tasa de interés del bono a un mes en un punto porcentual, generará una variación en la probabilidad de ganar o perder de -5.99%, de acuerdo al APE. Por el contrario, un movimiento en el tipo de cambio de un peso, provocará un cambio en la probabilidad de ganar de 2.191% utilizando al PEA. Para finalizar, un movimiento equivalente a un peso en el precio de cierre de BIMBOA provocará una variación de 2.586% en la probabilidad de ganar o perder en dicha acción. Pero una variación en el tipo de cambio en la misma magnitud conllevará una reducción o crecimiento en la probabilidad de ganar en -1.737% usando el concepto PEA.

Conclusiones.

Tal como se puede observar en los resultados, el tipo de cambio y la tasa del bono si afectan directamente al desempeño de la acción, medido con el rendimiento. Ahora bien, el siguiente paso sería incorporar al modelo más variables, sean internas o externas, que constituyan a priori fuentes de riesgo, como el ratio deuda/capital, la beta de las empresas, los índices ROA y ROE, el índice nacional de precios al consumidor, nivel de desempleo, etc. Sabemos que la periodicidad es un factor clave, pero atendiendo a lo comentado en el apartado “Definición de muestra” los datos diarios provocan una mayor variabilidad, por lo que una mejor estimación podría lograrse con datos mensuales.

Creemos que las estimaciones son buenas, pero sesgar la modelación a cuatro empresas, aunque representativas, genera un problema a la hora de generalizar los resultados presentados a las aproximadamente 143 empresas en la BMV. Por tanto consideramos que el modelo debería sectorizarse por actividades productivas (minería, comunicaciones, transporte, alimentos, etc.). Y posteriormente aplicarse a otras bolsas teniendo en cuenta las ideas anteriores.

Bibliografía.

- De Sousa S. Fernando (2013). Modelos de Valoración de activos financieros (CAPM) y teoría de valoración por arbitraje (APT): Un test empírico. Universidade do Estado de Minas Gerais. Brasil.

- Francisco C. Berenice, Francisco P. Abel (2016). Aplicación de la Teoría del Teorema Modigliani-Miller. Universidad Nacional Pedro Henríquez Ureña. República Dominicana.
- Puig, X. (2016). Split o desdoblamiento de acciones. Diccionario de Economía. El Economista Información extraída de: <https://www.eleconomista.es/diccionario-de-economia/split>
- BBVA. (2019) ¿Qué es el riesgo financiero? Educación Financiera-Banca Comercial. Información extraída de: <https://www.bbva.com/es/finanzas-para-todos-el-riesgo-financiero-y-sus-tipos/>
- Gil V, Gerardo. (2015). Ganancias de Capital. Enciclopedia Jurídica Online. Información extraída de: <https://mexico.leyderecho.org/ganancias-de-capital/>